

UNITED STATES LETTERS PATENT APPLICATION
FOR:

METHOD AND SYSTEM FOR RECOGNIZING DATE

INVENTOR:

William Patrick Tunney

Prepared by:

KENYON & KENYON
One Broadway
New York, NY 10004
(212) 425-7200

METHOD AND SYSTEM FOR RECOGNIZING DATE

BACKGROUND

[001] Handwriting recognition software has made it possible to digitally capture 5 handwriting and transform it into digital characters using an input capture device and a computer. The capture device may be a flat panel device that allows a user to enter normal handwritten notes onto a piece of paper attached to the capture device while information about the coordinates of the pen strokes is digitally recorded by the capture device. The capture device can later upload the digitally recorded handwritten notes 10 into a computer where an uploading program receives and stores the handwriting strokes in memory, resulting in two copies of a document, namely the original handwritten version and a second, digitally encoded version.

[002] Digital handwriting capture is useful when data must be entered into a 15 computer program for processing, but original handwritten copies must be retained for legal or verification purposes. In these instances, it would be helpful to have handwriting automatically transformed into digital characters and transferred to a computer program without manual data entry. This may be achieved by placing a printed paper form with clearly defined input fields on a capture device, digitally 20 capturing the handwritten notes in these input fields on the capture device, and uploading the digital notes to the computer. A recognition program may then interpret the digitally recorded handwritten notes within these input fields and transform them into a digitally encoded representation, which can be automatically entered into the computer program in the same manner as if the notes were manually entered via a 25 keyboard.

[003] However, the problem with these systems is that handwriting recognition software is generally not 100% accurate. Conventionally, an accuracy of approximately 30 90% can be reached, but only after a user has gone through a series of arduous user-specific training sessions that allow the recognition software to adjust to the style of a particular user. So, time is lost by having to train the software. But, even so, there may

still be multiple transformation errors because of the margin for error in current handwriting recognition software.

[004] Errors are particularly prevalent when a user enters calendar date information in numerical form. For example, if an input field is provided for date entry and no format is given, a user may enter the date in one of many formats. December 2, 1975 may appear as 12.02.1975 or 12/02/75 in the U.S. and 02.12.1975 or 02/12/75 in Europe, for example. The ambiguous nature of the date field results in difficulty in resolving handwritten input as a date understood by the computer program that will later use this information. Additionally, even if separate fields are provided for day, month, and year and the expected format is clearly marked on the printed form, users may still unintentionally enter the date in an incorrect format if they ignore this additional instruction.

[005] Finally, even if users correctly follow the format required, other difficulties may arise when resolving individual numbers. Different cultures may write numbers differently. For example, the handwritten German 1 closely resembles a 7 and is interpreted as such by handwriting recognition software developed in the U.S. Or handwriting recognition engines may lack robustness in interpreting numbers. For example, handwriting recognition engines that are trained on a specific user's input may not correctly interpret hastily written numbers by the specific user and can not at all interpret numbers from a different user. In some cases, the number 8 could easily be interpreted as the letter B.

[006] Some systems have tried to solve these problems with date recognition by providing graphical user interfaces through which a user may select date information from pull-down menus. In these systems, a more complex input/output device than the capture device must be used to display the graphical user interfaces. Such a device could be expensive and too bulky to carry, particularly for field surveys, field inventory, etc., for which the capture device is ideally suited.

[007] Accordingly, there is a need in the art for a simple and natural way to improve the recognition accuracy of calendar date information entered by a user onto printed paper forms attached to capture devices independent of the user who inputs the information.

5

SUMMARY OF THE INVENTION

[008] Embodiments of the present invention provide a simple and natural method to improve recognition of calendar date information. These embodiments provide date information by making marks on a paper data form corresponding to the 10 month, day, and year, while a capture device captures when and where on the paper data form the marks were made. A method includes a processor receiving capture data from the capture device, where the capture data is captured simultaneously with writing made on the paper form. The method further includes the processor comparing the capture data with a plurality of unique positions stored in memory in association with a 15 plurality of calendar dates printed on the paper form, retrieving from memory the calendar date associated with the unique position that matches the capture data, and storing in memory the retrieved calendar date as the writing made on the paper form. The capture data is advantageously generated by simply using a piece of paper and the capture device without having to rely on more complex, bulky devices with graphical 20 user interfaces.

[009] Embodiments of the present invention also provide a system through 25 which date information may be recognized. The system may include a memory and a processor for receiving capture data corresponding to a set of marks made on a paper data form attached to a capture device and mapping the capture data to a date.

BRIEF DESCRIPTION OF DRAWINGS

[0010] FIG. 1 is an exemplary computer network used to recognize the date according to embodiments of the present invention.

30

[0011] FIG. 2 is an exemplary computer used to recognize the date according to 35 embodiments of the present invention.

[0012] FIG. 3 is an exemplary paper data form that includes date information to be filled out according to an embodiment of the present invention.

5 **[0013]** FIG. 4 is an exemplary data capture format according to an embodiment of the present invention.

[0014] FIG. 5 is a flowchart of an embodiment of a method according to the present invention.

10

DETAILED DESCRIPTION

[0015] Embodiments of the present invention provide a method and system for recognizing calendar date information from a paper data form attached to a capture device. Date information on the paper may be in the form of a set of boxes representing the months of the year, the days of the month, and a current span of years. A user may simply fill in the boxes corresponding to the correct date. The capture device may digitally capture the pen strokes the user makes when filling in the boxes and upload this capture data to a computer for date recognition according to embodiments of the present invention. Exemplary applications of these embodiments include field surveys, field inventory, and other applications where paper forms are the predominant way data is recorded and device portability and ease of use are preferable.

20 **[0016]** In embodiments of the present invention, the computer's processor may receive the capture data from a capture device to which a paper data form was previously attached. The capture data format may be a time ordered sequence of (x,y) coordinates, indicating when and where on the data form a set of marks (or pen strokes) was made to fill in the correct date. This capture data may then be compared to a set of predefined coordinates stored in the computer's memory which indicates the expected positions of each month, day, and year on the data form. The predefined data that matches the capture data may be determined and the corresponding month, day, and year stored in memory for later use; hence, the date is recognized. In an

alternate embodiment, the capture data format may be a sequence of vectors (x, y, t) or any format that appropriately represents the user's pen strokes.

5 [0017] Instead of having to rely on handwritten numbers or a variety of date formats, embodiments of the present invention may use the capture data of the filled in boxes to determine the date. Accordingly, these embodiments advantageously provide a simple and natural way to accurately recognize data information. Additionally, problematic handwriting recognition need not be done, resulting in reduced data ambiguity and improved speed of data entry.

10

[0018] FIG. 1 shows an embodiment of an exemplary network that may be used to implement embodiments of the present invention. The exemplary network system 100 may include, but is not limited to, a computer network 110, computers 120-1 through 120-C, where C is an integer, capture devices 160-1 through 160-B used by 15 users 170-1 through 170-B, where B is an integer, to input date information, a server 140, and a database 150 storing various date coordinates used by the computers. These components may be linked to the network 110 via network links 115. The network 110 may be a LAN, WAN, Internet, or any like structure capable of connecting 20 components and transmitting data. The network links 115 may include physical wiring, wireless connections, or any like transmission configuration capable of transmitting data. Alternatively, a capture device 160 may be directly linked via a wireless link 117, a COM cable 119, or any like connector, to a computer 120.

25 [0019] The capture device used in embodiments of the present invention may include a portable input device whose appearance and operation resembles that of a traditional clipboard. The capture device may include a flat panel onto which a piece of paper may be attached and pens used to write on the paper thereby entering data to the capture device. The paper generally replaces a graphical user interface that is included in most input devices. So, typically, the capture device does not include a 30 graphical user interface. The pen strokes made on the paper may be stored in memory on the capture device for later uploading to a computer via a modem, cable, or other

transmission device in communication with a port of the capture device. An example of the capture device is the CrossPad™ manufactured by IBM.

[0020] In an embodiment, the capture device may include software for
5 interacting with a user and for uploading capture data to the computer. The capture device may include a series of built-in buttons that may be configured to initiate given commands. For example, capture data may be uploaded to the computer via the wireless link, COM cable, or the like, by the user pressing some of the buttons to initiate the upload process. After the upload completes, the user may delete the capture data
10 from the capture device. The capture device may include a small text-based display to show short text messages to the user.

[0021] In an alternate embodiment, the capture device may include local intelligence for performing recognition and uploading the recognized data to the
15 computer for further processing.

[0022] Since digital handwriting capture is not limited to physical flat panel devices, in another alternative embodiment the capture device may include electronic reusable paper, for example. Electronic reusable paper is designed to have the look and
20 feel of normal paper, except that it contains tiny sensor network technologies that provide digital display and capture of handwritten notes. Similar to a flat panel device, data can be captured, except that in the case of electronic reusable paper that data is collected and stored by the paper itself. Data collection from electronic reusable paper may be implemented in many ways, including attaching the paper to a clipboard
25 containing the electronics required to retrieve data from the electronic reusable paper and forwarding the data obtained using standard methods. An example of electronic reusable paper is SmartPaper manufactured by Gyron LLC.

[0023] FIG. 2 is a block diagram of an exemplary computer that can implement
30 embodiments of the present invention. The computer 200 may receive capture data from the capture device according to embodiments of the present invention. The computer 200 may include, but is not limited to, a processor 220 provided in

communication with a system memory module 230, a storage device 240, and an I/O device 250. The processor 220 may perform data recognition with the capture data received from the capture device. The memory 230 may store program instructions to be executed by the processor 220 and also may store variable data generated pursuant
5 to program execution. In practice, the memory 230 may be a memory system including one or more electrical, magnetic, or optical memory devices. The I/O device 250 may include a docking station for interface to the capture device 160 to receive the capture data and transmit any other appropriate data between the capture device 160 and the computer 200.

10

[0024] In embodiments of the present invention, a paper form may have printed thereon date information, including 12 months, 31 days, and a current span of years. Each possible value for month, day, and year may occupy a unique spatial position on the paper form. When a user marks a date on the form by selecting a month, day, and
15 year, the positions of the user's marks are unique to that date. These positions may be captured by the capture device and later uploaded to a computer for processing. Hence, the computer may determine the calendar date based on these unique positions.

20

[0025] FIG. 3 is an example of a paper data form in which the date information is printed in unique positions as described. In this example, the data form 300 may include, but is not limited to, date information comprising a first group of boxes 310 listing the 12 months in a year, a second group of boxes 320 listing 31 days in a month, and a third group of boxes 330 listing a current span of years. Each box 310, 320, 330 occupies a unique position on the data form 300, which, when marked by the user, may
25 be determined based on that position of the mark. FIG. 4 will further illustrate this.

30

[0026] The data form 300 may also include time information comprising a clock 340 and a couple of boxes 350 listing "AM" and "PM." The data form 300 may also include a survey 360 to be filled out and the identification 370 of the data form.

[0027] The data form 300 may be attached to the capture device 160 and the correct date boxes filled in from the first, second, and third groups of boxes 310, 320,

330. The coordinates of the marks made to fill in the boxes may be recorded on the capture device 160 and later uploaded to the computer 120 for processing according to embodiments of the present invention. The boxes may be filled in by shading the entire box or making a checkmark, an "X", or any discernible indicator that the box has been
5 selected.

[0028] To accurately capture the correct coordinates of the date information, the data form 300 should first be properly oriented on the capture device 160 and firmly attached to prevent movement. In one embodiment, a border or like markers may be
10 printed on the face of the capture device indicating where the data form 300 should be attached. Alternatively, the data form 300 may have printed in each corner a hash mark or like markers. A user first would write on the paper form at the hash marks prior to marking the form with date information. The coordinates of these hash marks may be captured and uploaded to the computer where used as reference points for the
15 coordinates of the date information.

[0029] Different data forms 300 may have the date information 310, 320, 330 in different positions and/or formats on the form. Hence, the identification 370 of the data form may be uploaded to the computer 120 so that the computer 120 may retrieve the
20 appropriate predefined coordinates. In one embodiment, the form identification 370 may have a uniquely positioned box associated with it to be filled in by the user. The coordinates of the filled-in box may indicate to the computer 120 which data form is being used.

25 **[0030]** It is to be understood that the layout of the date information on the form is not limited to groups of boxes, as shown in FIG. 3. The date information may have any layout that allows the user to mark a date on a form without having to rely on handwriting recognition software. For example, in an alternate data form, the day information 320 may resemble a page of a calendar and a day in the calendar filled in.
30

[0031] FIG. 4 illustrates an example of the capture data format that may be used in embodiments of the present invention. In this example, the user filled in, using "X"

marks, the appropriate boxes 310, 320, 330 on the paper data form to indicate the date June 5, 2003. The capture device digitally captured these "X" marks as time ordered coordinates 440, 450, 460. Here, (a1,b1) are the unique coordinates for the "Jun" box, (a2,b2) are the unique coordinates for the "5" box, and (a3,b3) are the unique
5 coordinates for the "2003" box. The user checked the boxes in the order of "Jun," "5," "2003." Hence, the corresponding coordinates were uploaded to the computer in that order, as illustrated by 470. The processor 220 may map the unique coordinates to the month, day, and year, i.e., Jun 5, 2003, to which they belong. The processor 220 may further use the ordering as indication of when the marks were made, i.e., relative to
10 each other.

[0032] FIG. 5 is a flowchart of an embodiment of a method for recognizing date information according to the present invention. The processor 220 may receive (505) capture data from the capture device 160. As stated previously, the capture data may
15 include, but is not limited to, a time-ordered set of coordinates indicating when and where on a paper data form a set of marks was made to fill in the correct date. The processor 220 may then compare (510) the set of coordinates with a set of predefined coordinates in memory 230 or storage 240 to find a match for the month, day, and year. The predefined coordinates may define the expected unique positions of marks on the
20 paper data form for each month, day, and year.

[0033] In a system where a variety of paper data forms may be used, the processor 220 may also receive the form identification from the capture device 160. The form identification may be indicated by a set of coordinates, vectors, etc., indicating
25 where on the paper data form the form mark was made. Prior to retrieving the predefined coordinates, the processor 220 may compare the form identification coordinates with predefined identification coordinates to find a match. The processor 220 may then determine the predefined coordinates in memory 230 or storage 240 based on the form identification and compare (510) the captured set of coordinates with
30 these determined predefined coordinates.

[0034] When a user incorrectly fills in the date information and then corrects the mistake, the capture device 160 records multiple coordinates within the same group of boxes for the month and/or day and/or year. For example, if the user mistakenly selected the month as October, scratched it out, and then correctly selected the month as November, there would be three discrete sets of captured coordinates in the group of month boxes—a first set made in selecting October, a second set in scratching out October, and a third set in selecting November—where each set would be spaced out in time. In this case, the processor 220 will have incorrectly received multiple coordinates. However, as previously described, the capture device 160 captures the time when the mark was made, either implicitly, in the ordering of the sequence of (x,y) coordinates, or explicitly, in the vector coordinates (x,y,t), for example. So, the processor 220 solves this problem of multiple coordinates using the time data.

[0035] If the processor 220 receives (520) multiple coordinates within a group, i.e., finds multiple matches with the predefined coordinates, the processor 220 may determine (525) the set of coordinates in the group that was captured last by the capture device 160. This last set corresponds to the correct marks made by the user on the data form and, hence, the correct predefined coordinates.

[0036] Next, the processor 220 may retrieve (530) from memory or storage the date associated with the predefined coordinates that match the captured set of coordinates. The processor 220 may then store (540) the month, day, and year as the date the user marked on the form.

[0037] The processor 220 may alternatively retrieve the predefined coordinates from memory or storage, one at a time or together, prior to the comparison with the captured set of coordinates and then store the month, day, and year that match the captured set of coordinates as the date the user marked on the form.

[0038] In an alternate embodiment, the capture device 160 may perform both the data capture and the date determination. After which, the capture device 160 may upload the date to the computer 120 for further use or storage.

[0039] Embodiments of the present invention may be implemented using any type of computer, such as a general-purpose microprocessor, programmed according to the teachings of the embodiments. The embodiments of the present invention thus also

5 includes a machine readable medium, which may include instructions used to program a processor to perform a method according to the embodiments of the present invention. This medium may include, but is not limited to, any type of disk including floppy disk, optical disk, and CD-ROMs.

10 [0040] It may be understood that the structure of the software used to implement the embodiments of the invention may take any desired form, such as a single or multiple programs. It may be further understood that the method of an embodiment of the present invention may be implemented by software, hardware, or a combination thereof.

15 [0041] The above is a detailed discussion of the preferred embodiments of the invention. The full scope of the invention to which applicants are entitled is defined by the claims hereinafter. It is intended that the scope of the claims may cover other embodiments than those described above and their equivalents.